Rules and Heuristics during Sentence Comprehension: Evidence from a Dual-task Brain Potential Study

Manuel Martín-Loeches[1,2], Annekathrin Schacht[3], Pilar Casado[1], Annette Hohlfeld[3], Rasha Abdel Rahman[3], and Werner Sommer[3]

Abstract

■ Whether syntactic and semantic processes during sentence comprehension follow strict sets of rules or succumb to context-dependent heuristics was studied by recording event-related potentials in a dual-task design. In one condition, sentence-extraneous acoustic material was either semantically congruent or incongruent relative to an adjective in the visually presented sentence, the latter being either semantically correct or incorrect within the sentence context. Homologous syntactic (gender) manipulations were performed in another condition. Syntactic processing within the sentence appeared to be blind to the syntactic content of the second task. In contrast, semantically incongruous material of the second task induced fluctuations typically associated with the detection of within-sentence semantic anomalies (N400) even in semantically correct sentences. Subtle but extant differences in topography between this N400 and that obtained with within-sentence semantic violations add to recent proposals of separate semantic subsystems differing in their specificity for sentence structure and computational procedures. Semantically incongruous material of the second task also influenced later stages of the processing of semantically incorrect adjectives (P600 component), which are traditionally assumed to pertain to the syntactic domain. This result is discussed in the light of current proposals of a third combinatorial stream in sentence comprehension. ■

INTRODUCTION

During sentence processing, the meaning of individual words and the syntactic structure of the sentence must be combined in order to achieve full comprehension. It is a matter of debate in psycholinguistics exactly how conceptual/semantic and syntactic information are implemented within the sentence processing machinery. One kind of model holds that separable, independent, and at least partly sequential processes construct distinct syntactic and semantic representations (Ferreira & Clifton, 1986; Berwick & Weinberg, 1984). On the other side, fully interactive models suggest that syntactic and semantic constraints interact directly and simultaneously with each other at the message-level representation of the input (McClelland, St John, & Taraban, 1989; Marslen-Wilson & Tyler, 1987; Johnson-Laird, 1983). Other intermediate proposals differ in the degree of independence and prevalence ascribed to conceptual/semantic and syntactic information (e.g., Kim & Osterhout, 2005; Trueswell, Tannenhaus, & Garnsey, 1994; Frazier, 1987).

In order to elucidate the presence and nature of interactions between the syntactic and the semantic system, a number of open questions need to be resolved. In general, the semantic and syntactic streams appear to exhibit qualitatively different properties. The syntactic stream has always been considered as strictly algorithmic, following a finite list of well-defined instructions governing how words and other lexical elements combine to form phrases and sentences (Friederici & Weissenborn, 2007). Accordingly, syntactic processes assign thematic roles on the basis of morphosyntactic constraints. In contrast, the semantic system appears to be a heuristic combinatorial system. In this conception, sentences are treated essentially as unordered lists of words that combine lexical items based on plausibility according to our variable and flexible world-knowledge (Jackendoff, 2007; Vissers, Chwilla, & Kolk, 2007). Several recent models emphasize the heuristic nature of semantic information as the main determinant for sentence comprehension (Ferreira, 2003; Townsend & Bever, 2001).

Alas, the views of the semantic and the syntactic systems given above are far from being exhaustive (e.g., MacDonald & Seidenberg, 2006). As a case in point, Kuperberg (2007) suggests that semantic information during sentence processing should be subdivided into associative memory-based semantic relationships on the one hand and semantic–thematic relationships that have implications for the syntactic structure on the

[1]UCM-ISCIII, Madrid, Spain, [2]Complutense University of Madrid, Spain, [3]Humboldt-University at Berlin, Germany

other hand. Whereas the heuristic nature of the former semantic constituent is obvious, the same cannot be said for semantic–thematic relationships. Indeed, a recent fMRI study by Humphries, Binder, Medler, and Liebenthal (2007) indicates anatomically independent subsystems devoted to processing semantic sentence content and the meaning of individual words. Furthermore, it is also possible that the syntactic system is not totally algorithmic because several authors have stressed the relevance of some heuristics used in sentence comprehension, such as word order, that appear syntactic in nature (Jackendoff, 2007; Townsend & Bever, 2001).

The purpose of the present article is to further explore the algorithmic versus heuristic nature of the syntactic and semantic streams during sentence processing. A novel approach was taken by combining a dual-task paradigm with the recording of event-related brain potentials (ERPs). ERPs offer high temporal resolution, permitting measurements of electrical brain activities as language processing unfolds over time. Different ERP components seem to honor the distinction between the processing of syntactic and semantic information. When semantic variables are manipulated, the so-called N400 effect is the main finding (Kutas & Hillyard, 1980). The N400 effect is a negative-going component between roughly 250 and 550 msec, usually most pronounced over central and posterior electrode sites (Kutas & Besson, 1999). Typically, this component increases in amplitude with the difficulty of integrating words into their semantic context—be it a sentence or a preceding prime word (Chwilla, Brown, & Hagoort, 1995).

When a manipulated variable belongs to the syntactic domain, the main ERP effects are anterior negativities and posterior positivities. Anterior negativities have been typically labeled as left anterior negativity (LAN), resembling the N400 in latency, or ELAN (early LAN), appearing as early as 100 to 200 msec. Word category violations are the variations most frequently associated with ELAN (e.g., Friederici & Mecklinger, 1996), whereas other grammatical anomalies, including morphosyntactic violations (e.g., Coulson, King, & Kutas, 1998), usually evoke a LAN. Both anterior negativities may reflect highly automatic first-pass parsing processes, the detection of a morphosyntactic mismatch, and/or the inability to assign the incoming word to the current phrase structure (Friederici, 2002).

Finally, a late positive-going component with a parietal maximum, labeled P600, has typically been considered as a syntax-related ERP fluctuation because it is elicited by syntactic violations (e.g., Osterhout & Holcomb, 1992). However, a P600 also appears in structurally ambiguous or garden path sentences (Frisch, Schlesewsky, Saddy, & Alpermann, 2002). Therefore, it has been suggested that the P600 indicates increased syntactic processing costs due to necessary revisions and reanalyses of structural mismatches, possibly also reflecting subsequent repair processes (Münte, Heinze, Matzke, Wieringa, & Johannes, 1998). Although the P600 is predominantly viewed as a syntax-related ERP component (Kutas, Van Petten, & Kluender, 2006), the occasional observation of P600 deflections to purely semantic violations has motivated recent alternative proposals. According to one of these proposals, the P600 might reflect the activity of a combinatorial system that integrates both semantic and syntactic information. This system would still be syntactic in nature because its main function would be the assignment of thematic roles (Kuperberg, 2007). Another suggestion is that the P600 reflects a domain-general monitoring mechanism (Kolk & Chwilla, 2007).

Overall, a review of the ERP literature appears to indicate that syntax-related components (ELAN, LAN, and P600) only appear in the frame of sentence processing (Barber & Carreiras, 2005), whereas the semantic-related N400 component has been reported with similar latency, amplitude, and topography even for isolated word pairs (Kutas, 1993) outside of a sentence context. Already, these data indicate an algorithmic versus heuristic nature of syntactic and semantic processing streams, respectively. However, as the available evidence is limited, further research is warranted.

A useful experimental approach to distinguish between algorithmic and heuristic processes is the dual-task paradigm. In this paradigm, a task of interest is combined with a second task for stimuli that are delivered at defined points in time relative to the first task, allowing numerous inferences. In psycholinguistics, dual tasks have been used in the study of phonological (Reynolds & Besner, 2006) and lexical processing (Allen et al., 2002; McCann, Remington, & Van Selst, 2000), as well as language production (Ferreira & Pashler, 2002). In the present study, we used a dual-task paradigm to assess the question of whether syntactic and semantic processes differ in their algorithmic and heuristic processing properties, respectively. To this end, written sentences, shown word by word, were read while sentence-extraneous spoken material had to be stored and held in working memory. By recording ERPs, we investigated whether syntactic and semantic processes might be differentially susceptible to these extraneous influences.

In particular, the written sentences (Task 1) could be correct or incorrect from either the syntactic or the semantic point of view. Both kinds of violation occurred in a particular word of the sentence. Relative to correct versions, syntactic violations were expected to yield anterior negativities and P600, whereas semantic violations were expected to yield an N400 and, possibly, also a P600. Each sentence had to be judged for correctness following its presentation. As an example, the sentence *Los enemigos*$_{\text{[masc.]}}$ *agresivos*$_{\text{[masc.]}}$ *luchan* (nonliterally: *The enemies*$_{[masc.]}$ *aggressive*$_{[masc.]}$ *fight*) could be violated syntactically by modifying the gender of the adjective (*agresivas*$_{\text{[fem.]}}$), and semantically by replacing the correct adjective by an inappropriate one (*opacos*$_{\text{[masc.]}}$ = *opaque*$_{\text{[masc.]}}$).

Shortly before the violation within the written sentence, a spoken word was presented. Participants had to keep this word in working memory (Task 2), as they were to repeat it after giving correctness judgments about the sentence. The spoken word could exhibit various features that might differentially modulate the ERP fluctuations related to syntactic or semantic violations within the sentence. Specifically, a syntactic violation of morphosyntactic type (gender agreement violation in Spanish) within the written sentence was preceded by a spoken word of feminine or masculine form that either matched (*opacas*$_{[fem.]}$ = *opaque*$_{[fem.]}$) or mismatched (*opacos*$_{[masc.]}$ = *opaque*$_{[masc.]}$) the violation in the sentence. A semantic violation, on the other hand, was preceded by a spoken word that semantically matched (*coléricas*$_{[fem.]}$ = *furious*$_{[fem.]}$) or mismatched (*veladas*$_{[fem.]}$ = *fogged*$_{[fem.]}$) the violation in the sentence. Correct sentence material was also used, relative to which the spoken word could match or mismatch syntactically or semantically.

If sentence-extraneous acoustic material (the spoken words) influences either syntactic or semantic processing, as measured by ERPs, the affected process could be considered heuristic rather than algorithmic. In general, and according to most of the literature, a differential influence can be expected. Syntactic processing should be unaffected by syntactic material not pertaining to the sentence, whereas semantic processing should be amenable to sentence-extraneous semantic material. Specifically, if the semantic system is, in fact, heuristic in nature, an N400 should be obtained even to correct sentential material if accompanied by semantically incongruous sentence-extraneous material. In its most extreme case, the N400 elicited by sentence-extraneous incongruence might be identical to that elicited by within-sentence violations. If, on the other hand, the syntactic system is algorithmic as suggested by many, no LAN should be obtained for correct sentence material when accompanied by syntactically incongruent sentence-extraneous material. Likewise, the LAN obtained for syntactically incorrect sentence material should not be affected by sentence-extraneous syntactic incongruence.

The pattern obtained with the P600 should have implications for several controversial issues. If the P600 obtained to sentence-internal syntactic violation is the consequence of an algorithmic syntactic process, it should be unaffected by sentence-extraneous syntactic variations. Interestingly, there are two different options for the P600 possibly elicited by semantic within-sentence violations. On the one hand, the P600 could reflect essentially syntactic processes (thematic role assignment), and might therefore be immune to semantic incongruence of sentence-extraneous material. Alternatively, if this component reflects the activity of a third, domain-general stream, the P600 to semantic within-sentence violations might be affected by sentence-extraneous semantic material.

METHODS

Participants

After eliminating datasets of four participants with bad recordings, data from 32 native Spanish speakers (29 women), ranging in age from 18 to 23 years (M = 19.1 years), were analyzed. All had normal or corrected-to-normal vision and were right-handed, with average handedness scores (Oldfield, 1971) of +81, ranging from +40 to +100. The study was performed in accordance with the Declaration of Helsinki, and was approved by the ethics committee of the Center for Human Evolution and Behavior, UCM-ISCIII, Madrid, Spain. Participants gave their informed consent prior to the study and received reimbursement thereafter.

Materials

Task 1: Sentence Processing

In Task 1, the set of experimental items consisted of 160 Spanish *correct sentences* of the structure, [Det]–[N]–[Adj]–[V] (determiner–noun–adjective–verb). In these sentences, all nouns and adjectives are marked for gender. In addition to the correct version of each sentence, two unacceptable versions were created. One contained a *semantic violation* due to an unacceptable combination of noun and adjective. The second version contained a *syntactic violation* of the gender agreement between the noun and the adjective by modifying the latter. In all three versions of the sentences, the critical adjectives were of comparable frequency (20 per million) according to the Lexico Informatizado del Español (LEXESP; Sebastián, 2000) and length (Ms = 7.6 letters for correct and syntactically anomalous adjectives and 7.5 for semantically anomalous adjectives). Furthermore, a set of 160 filler sentences was constructed. Half of them were short fillers and followed the same structure as the experimental materials but the adjectives were omitted. For the remaining—long—fillers, a complement was appended to the structure of the experimental sentences. Half of both short and long fillers were unacceptable sentences with syntactic or semantic violations—depending on condition—either in the verb or in the complement, for short and long fillers, respectively. All stimuli in Task 1 were presented white-on-black on a computer monitor, controlled by Presentation Software. Participants' eyes were about 65 cm away from the monitor, yielding viewing angles between 0.7° and 1.3° in height and 1.1° to 6° in width.

Task 2: Memory Task

For Task 2, a set of 480 spoken adjectives was constructed according to the following principles: 160 adjectives were semantically related to both the noun and the adjective in a given correct experimental sentence of Task 1. In contrast, a second group of 160 adjectives had no semantic relationship with either noun or adjective in the corresponding correct experimental sentences of Task 1. A final group of 160 adjectives was semantically unrelated to the nouns and adjectives in the correct experimental sentences of Task 1, and to all the adjectives of the second group. These principles were applied in order to achieve appropriate combinations of Task 1 versus Task 2 adjectives, as will be described in the Procedure section. Adjectives for the filler sentences were constructed following similar principles as for the experimental sentences. All adjectives within Task 2 were comparable in intensity and voice of speaker and were presented by means of loudspeakers located in front of the participants. Overall intensity of the acoustic adjectives was adjusted to a comfortable level for each participant.

Procedure

Participants performed two tasks simultaneously. In Task 1, participants judged each sentence for acceptability by pressing a button as soon as they detected an unacceptable word, or just after the last word for correct sentences. Correctness judgments were given with the index fingers. The assignment of hand to response type was counterbalanced. All sentences began with a fixation cross of 500 msec duration and were presented word-by-word, with 300 msec duration per word and a 600-msec SOA, allowing 4300 msec between the end of the last word in a sentence and the appearance of the first word in the next sentence. The first word in each sentence began with a capital letter and the last word ended with a period.

From the pool of 160 correct experimental sentences and their semantically or syntactically unacceptable counterparts, two blocks each of stimulus material for the syntactic and semantic conditions were constructed. For the syntactic condition, participants were presented with two blocks of sentences, each containing 80 correct and 80 syntactically incorrect experimental sentences. Within a given block, none of the experimental sentences was repeated, and if a sentence appeared in its correct version in one block, its incorrect version was included in the other block. The same logic was applied to the construction of the two sentence blocks for the semantic condition. A total of eight versions of such blocks was constructed per condition, such that each sentence appeared once across groups of eight participants in either the syntactic or the semantic condition, in either its correct or incorrect version, or being preceded by a matching or mismatching spoken adjective in Task 2. In addition, each block contained the whole set of 160 filler sentences. In blocks for the syntactic condition, 80% of the violations in the fillers were syntactic (subject–verb person disagreement) and 20% were semantic (incongruous verb). The opposite was true in blocks for the semantic condition. Within a given block, all sentences were presented in random order. The order of conditions was counterbalanced across participants.

Task 2 required participants to keep the spoken adjective in mind and to repeat it after the end of the written sentence. A question mark appeared on the screen for 1 sec, starting 1.3 sec after the last word of the sentence, prompting the repetition of the spoken adjective. Spoken word duration was variable but not longer than 550 msec for adjectives co-occurring with the experimental visual sentences of Task 1. Spoken adjectives co-occurring with filler sentences could have longer durations. The onset of the spoken word was always synchronized to the onset of the noun in the visual sentence. A scheme of the structure of an experimental trial is represented in Figure 1.

In the semantic condition, spoken adjectives matched or mismatched semantically with adjectives of the experimental sentences, but always mismatched syntactically (in gender) with adjectives of the experimental sentences. The same principles applied to the syntactic condition so that both correct and incorrect adjectives in the experimental written sentences were preceded by a spoken adjective that matched or mismatched the written adjective in gender, but always mismatching semantically with adjectives of the experimental sentences. Examples are given below, with word-by-word translations into English and nonliteral interpretations. The critical word (adjective) in Task 1 is underlined; matching and mismatching of the spoken adjectives (Task 2) always

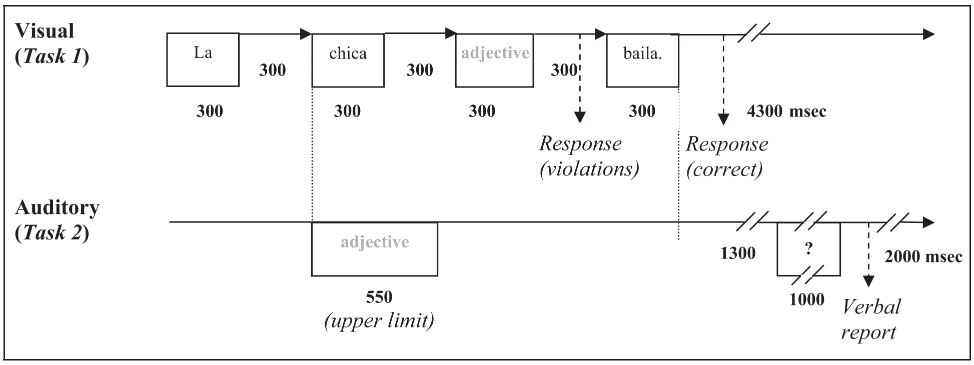

Figure 1. Schematic representation of the stimulation procedures. Two tasks were used simultaneously: Read a sentence presented word by word and judge for acceptability (Task 1), and hear a word, retain it, and say it aloud after the end of the sentence in write (Task 2).

refers to the written adjective in Task 1. More examples with word-by-word translations into English are provided in Appendix 1. As becomes clear from this scheme, the experimental manipulation of sentence correctness and matching between the spoken and the written adjective also has consequences for the relationship between the spoken adjective and the written noun. Whether this might affect the data is dealt with in the Discussion section.

Example for stimulus conditions are discussed below.

Semantic Condition

(Task 1) La chica$_{[fem.]}$ guapa$_{[fem.]}$ baila. (**Correct**)
The girl$_{[fem.]}$ beautiful$_{[fem.]}$ dances. (= *The beautiful girl dances*)

(Task 2) Bonito$_{[masc.]}$ (**Match**) Cuadrado$_{[masc.]}$ (**Mismatch**)
Pretty$_{[masc.]}$ Square$_{[masc.]}$

(Task 1) La chica$_{[fem.]}$ cuadrada$_{[fem.]}$ baila. (**Incorrect**)
The girl$_{[fem.]}$ square$_{[fem.]}$ dances. (= *The square girl dances*)

(Task 2) Redondo$_{[masc.]}$ (**Match**) Bonito$_{[masc.]}$ (**Mismatch**)
Round$_{[masc.]}$ Pretty$_{[masc.]}$

Syntactic Condition

(Task 1) La chica$_{[fem.]}$ guapa$_{[fem.]}$ baila. (**Correct**)
The girl$_{[fem.]}$ beautiful$_{[fem.]}$ dances. (= *The beautiful girl dances*)

(Task 2) Cuadrada$_{[fem.]}$ (**Match**) Cuadrado$_{[masc.]}$ (**Mismatch**)
Square$_{[fem.]}$ Square$_{[masc.]}$

(Task 1) La chica$_{[fem.]}$ guapo$_{[masc.]}$ baila. (**Incorrect**)
The girl$_{[fem.]}$ beautiful$_{[masc.]}$ dances. (= *The beautiful girl dances*)

(Task 2) Cuadrado$_{[masc.]}$ (**Match**) Cuadrada$_{[fem.]}$ (**Mismatch**)
Square$_{[masc.]}$ Square$_{[fem.]}$

Electrophysiological Recording and Analysis

The EEG was recorded from 27 tin electrodes mounted within an electrode cap (ElectroCap International; Eaton, OH); bandpass was 0.01 to 30 Hz, sampling was done at 250 Hz. All channels were referenced on-line to the right mastoid, and re-referenced off-line to the average of the left and right mastoids. Bipolar horizontal and vertical EOGs were recorded for artifact monitoring. All electrode impedances were kept below 3 kΩ.

The continuous EEG was segmented into 1600-msec epochs starting 200 msec before the onset of the noun in the experimental sentences. Artifacts were automatically rejected by eliminating epochs during which a range of ±100 μV was exceeded in any of the channels. Off-line, corrections for artifacts due to blinks, vertical, or horizontal eye movements were made using the method described by Gratton, Coles, and Donchin (1983). Based on visual inspection, all epochs were eliminating that still presented artifacts. Epochs with erroneous judgments or responses (correct sentences judged as unacceptable, incorrect sentences judged as acceptable, or incorrect verbal reports in Task 2) were also eliminated. Overall, the mean rejection rate was 34.4% (18.8% due to artifacts, the others due to response errors).

Separate average ERPs were calculated for epochs containing adjectives in the experimental sentences as a function of whether they were correct or not and preceded by a spoken Task 2 adjective matching or mismatching syntactically or semantically. Comparisons involved main effects of sentence correctness, main effects of matching, as well as their interaction, done separately for each condition (syntactic, semantic).

RESULTS

Sentence Correctness Judgments and Verbal Reports

Semantic Condition

Correct sentences were considered acceptable in 83.5% and 78.7% of the cases, respectively, when the spoken adjective of Task 2 semantically matched or mismatched the sentence adjective. Of the incorrect sentences containing semantic violations, 81.2% and 79.6%, respectively, were judged as unacceptable when the spoken adjective of Task 2 matched or mismatched the sentence adjective. ANOVA with factors acceptability (correct versus incorrect) and matching of spoken adjective (match versus mismatch) yielded a significant main effect of matching [$F(1, 31) = 8.47, p = .007$], but not of acceptability [$F(1, 31) = 0.05, p > .1$]. In addition, there was a trend for an interaction between these factors [$F(1, 31) = 3.29, p = .08$].

Reaction time (RT) data were skewed because responses to violations could be given immediately after their occurrence, whereas valid correctness judgments could be given only at the end of a sentence. As a result, mean RTs for correct sentences were 1001 and 1008 msec after adjective onset when the spoken adjectives of Task 2 semantically matched or mismatched, respectively. In contrast, RTs for incorrect sentences were 940 and 897 msec when spoken adjectives matched or mismatched, respectively. ANOVA yielded the expected main effect of acceptability [$F(1, 31) = 7.55, p < .01$], no effect of matching [$F(1, 31) = 2.42, p > .1$], but a significant interaction between these factors [$F(1, 31) = 5.77, p = .02$].

Overall, the performance data indicate that adjectives of Task 2, which semantically mismatch with adjectives in Task 1 sentences, induce more judgment errors in both correct and incorrect sentences as well as faster detections of semantic violations.

Syntactic Condition

In the syntactic condition, 82.7% of correct sentences were judged as acceptable when the spoken adjectives of Task 2 syntactically matched with the sentence's adjective and 80.7% were judged as acceptable when they mismatched. For syntactically incorrect sentences, 94.3% and 94.5% were judged as unacceptable. ANOVA yielded a significant main effect of acceptability [$F(1, 31) = 44.33$, $p < .0001$], but none for matching [$F(1, 31) = 1.37$, $p > .1$], nor was there an interaction [$F(1, 31) = 1.37$, $p > .10$].

Mean RTs in correct sentences were 995 and 994 msec when the spoken adjective of Task 2 syntactically matched or mismatched, respectively. In contrast, mean RTs were only 655 and 666 msec for incorrect sentences for matching and mismatching spoken adjectives, respectively. ANOVA yielded a significant main effect of acceptability [$F(1, 31) = 102.38$, $p < .0001$], but none for matching [$F(1, 31) = 0.24$, $p > .1$], nor was there an interaction [$F(1, 31) = 0.29$, $p > .1$].

According to these data, and at variance with the semantic condition, error rates were noticeably reduced for incorrect as compared to correct sentences, whereas the syntactic match or mismatch of the additional auditory material in Task 2 did not seem to induce any type of behavioral effect on sentence processing.

Memory Task Performance

Mean incorrect verbal reports in Task 2 were 14.33% ($SD = 7.2$) for the syntactic condition and 14.98% ($SD = 7.7$) for the semantic condition. The similarity in these error rates [$t(31) < 1$] indicates that differences in the degree of difficulty between the syntactic and the semantic conditions did not differentially affect the amount of resources employed for Task 2 in both conditions.

ERP Data

Preadjective Epochs

The data were first analyzed using the 200 msec preceding the noun as baseline. This analysis served two exploratory purposes. First, it was used to check for a

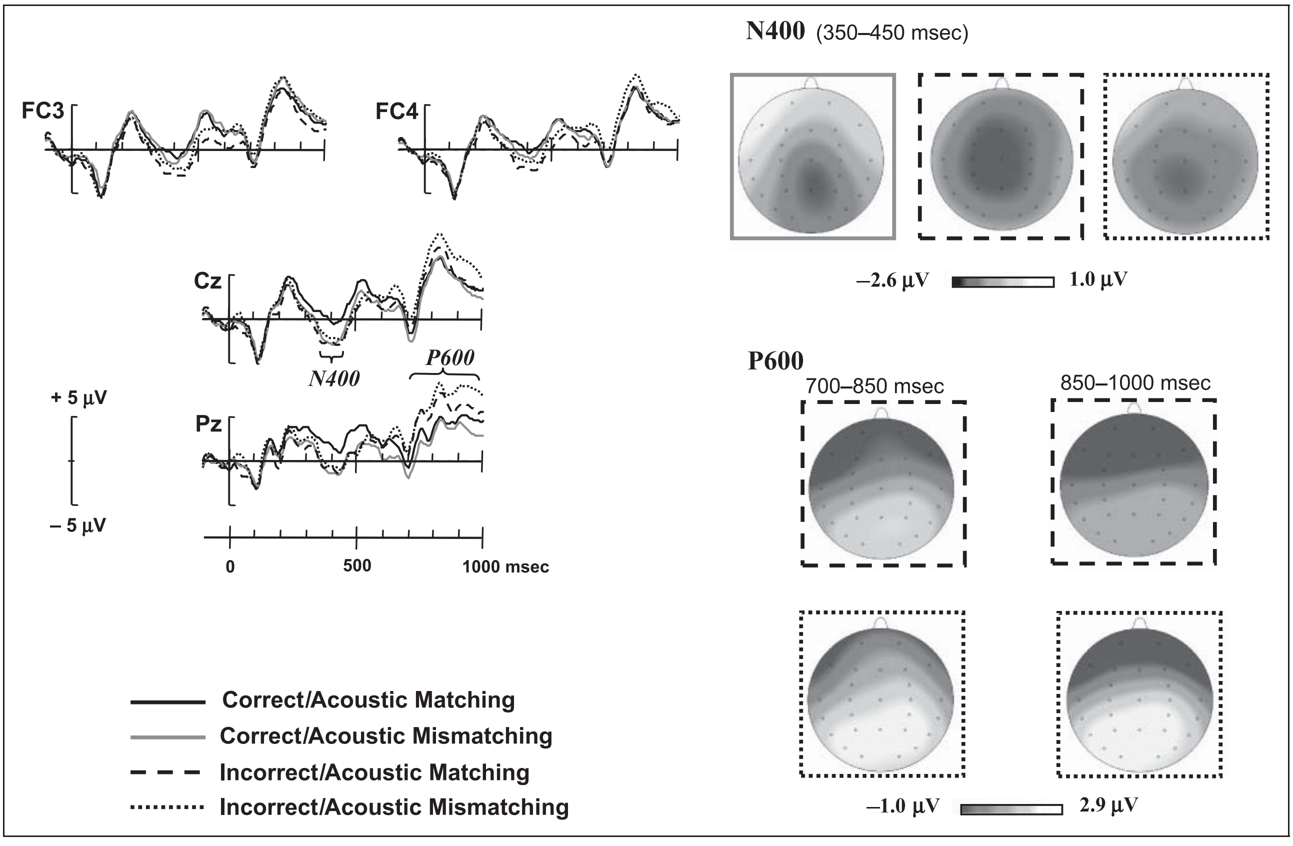

Figure 2. ERPs to semantically correct and incorrect adjectives, referred to a 100-msec prestimulus baseline, as a function of whether they were preceded by an acoustic stimulus matching or mismatching semantically with the adjective of the sentence. Left: ERP waveforms at a selection of electrodes. Right: Difference maps of the effects in the N400 and P600 time windows, interpolated with spherical splines (Perrin, Bertrand, & Echallier, 1989). The reference condition that is subtracted from the others is correct/matching.

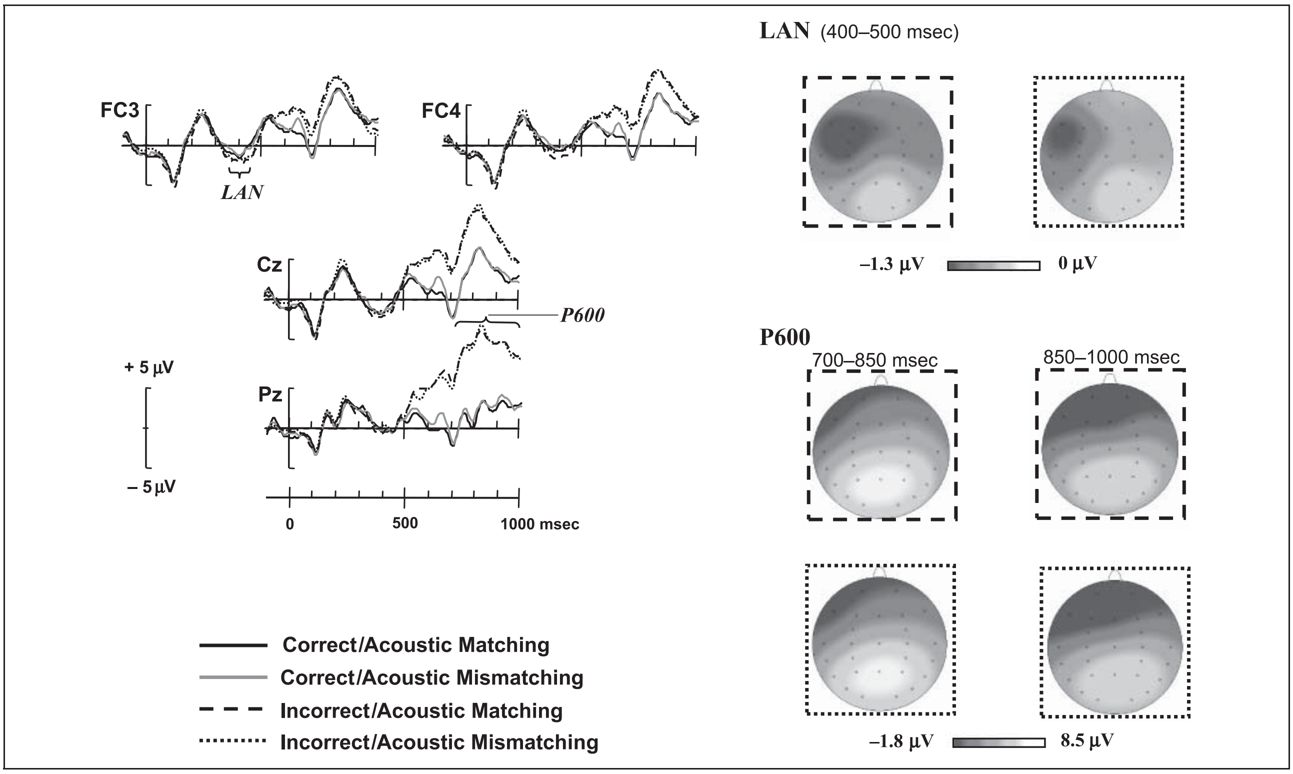

Figure 3. ERPs to syntactically correct and incorrect adjectives, referred to a 100-msec prestimulus baseline, as a function of whether they were preceded by an acoustic stimulus matching or mismatching syntactically (gender) with the adjective of the sentence. Left: ERP waveforms at a selection of electrodes. Right: Difference maps of the significant effects in the LAN and P600 time windows.

possible influence of the variation in acoustic parameters of Task 2 stimuli—namely, duration and intonation contours—on ERP fluctuations following the noun. In the absence of relevant effects within the interval between noun onset and adjective, a baseline immediately preceding the onset of the critical word (the adjective) appeared to be justified. Second, systematic ERP effects in the noun–adjective period would be informative about influences of the relationship between the spoken adjective on the semantic or syntactic processing of the written noun of the sentence.

Mean ERP amplitude of 12 consecutive 50-msec time windows following noun onset was submitted to repeated measures ANOVAs (Greenhouse–Geisser corrected) with factors sentence acceptability, matching of spoken adjective, and electrode site (27 levels). Note that the matching factor was defined relative to the upcoming written adjective in order to validate the preadjective baseline as unaffected by the variables of interest. A single significant effect of matching [$F(1, 31) = 4.43, p = .04$] was found in the 200–250 msec window of the syntactic condition. This effect appeared in an isolated window, relatively far from the onset of the adjective, and is meaningless relative to the noun (matching is relative to the adjective). Accordingly, the spoken adjective does not appear to noticeably influence the processing of the simultaneously appearing noun of the sentence, in line with previous reports that there are no priming effects when prime and target words are presented simultaneously (Holcomb & Anderson, 1993). Therefore, we confine the following report to ERP fluctuations following the onset of the written adjective, using a 100-msec preadjective baseline.

Adjective-related ERPs

Figure 2 summarizes main ERP results for the semantic condition, showing overlays of the ERP waveforms to the adjective in the four possible situations. Figure 3 displays the corresponding data for the syntactic condition. ERP mean amplitude measures were quantified in 17 consecutive 50-msec intervals starting 150 msec after the onset of the written adjective and lasting up to 1000 msec. The same ANOVAs were performed on these parameters as for the previous ERP analysis. Table 1 displays main significant results for the semantic and syntactic conditions.

Regarding the *semantic condition*, a clear N400 appeared for the two conditions with semantically incorrect adjectives, peaking at about 400 msec after stimulus onset and being most pronounced at Pz and Cz electrode sites. These two N400 deflections appeared to be very similar, regardless of whether the spoken adjective of Task 2 matched with the violating written adjective.

Table 1. Statistical Analyses for 17 Consecutive 50-msec Time Windows following the Onset of the Written Adjective

	df	*150–200*	*200–250*	*250–300*	*300–350*	*350–400*	*400–450*	*450–500*	*500–550*
Semantic Condition									
Acceptability	1, 31			11.71 (.002)	12.90 (.0001)	16.85 (.0001)	14.20 (.001)	15.8 (.0001)	16.92 (.0001)
Matching of Task 2	1, 31								
Electrode × Acceptability	26, 806								4.65 (.003)
Electrode × Matching of Task 2	26, 806				2.93 (.036)	5.03 (.003)	3.74 (.011)	7.61 (.0001)	4.73 (.004)
Acceptability × Matching of Task 2	1, 31					5.30 (.028)			
Electrode × Acceptability × Matching of Task 2	26, 806			3.3 (.021)	3.01 (.02)	2.82 (.039)			
Syntactic Condition									
Acceptability									
Matching of Task 2									
Electrode × Acceptability							2.69 (.03)	2.66 (.03)	5.86 (.001)
Electrode × Matching of Task 2									
Acceptability × Matching of Task 2									
Electrode × Acceptability × Matching of Task 2									

Only significant results are displayed.

Significant main effects of acceptability in the windows from 250 to 550 msec statistically support this impression. Remarkably, a similar N400 could also be seen for correct written adjectives if preceded by semantically incongruous spoken adjectives. This observation is statistically confirmed by significant main effects of matching and interactions of matching, acceptability, and electrode in the windows corresponding to the N400. Post hoc analyses performed on mean amplitudes at the Pz electrode between 350 and 450 msec after adjective onset corroborated that N400 amplitudes to correct adjectives preceded by a matching spoken adjective were significantly smaller as compared to each of the other three conditions [$3.43 \leq t(31) \leq 3.84$, *ps* $< .01$], whereas the three latter conditions did not differ among each other [$-0.6 \leq t(31) \leq -0.2$, *ps* $> .1$ (all *ps* Bonferroni corrected)].

As shown in the maps of Figure 2, the N400 for violations, particularly for the one preceded by matching spoken adjectives, displayed a wider distribution than that to correct adjectives preceded by mismatching spoken adjectives, covering also more frontal areas. In order to statistically compare the three N400 distributions, a profile analysis was performed with normalized data. Overall ANOVAs using the 350–450 msec window indicated that the three N400 components—the two for violations and the one for correct words preceded by mismatching spoken adjectives—did not significantly differ in topography [$F(52, 1612) = 1.2$, $p > .1$]. However, at least one of the pairwise comparisons—even when

550–600	600–650	650–700	700–750	750–800	800–850	850–900	900–950	950–1000
			10.92 (.002)	14.71 (.001)	15.7 (.0001)		8.01 (.008)	4.92 (.034)
4.4 (.003)	7.13 (.0001)	8.31 (.0001)	13.33 (.0001)	18.17 (.0001)	17.50 (.0001)	15.42 (.0001)	17.24 (.0001)	19.02 (.0001)
2.8 (.035)	2.83 (.036)							
	5.02 (.032)						4.27 (.047)	
						2.73 (.038)	3.06 (.034)	5.07 (.005)
22.55 (.0001)	41.33 (.0001)	45.10 (.0001)	69.61 (.0001)	105.70 (.0001)	64.72 (.0001)	41.53 (.0001)	18.90 (.0001)	7.50 (.01)
10.96 (.0001)	15.05 (.0001)	25.12 (.0001)	38.60 (.0001)	74.77 (.0001)	73.50 (.0001)	44.42 (.0001)	49.67 (.0001)	56.71 (.0001)

Bonferroni-corrected—revealed a significant difference in topography, namely, between the N400 to correct adjectives preceded by mismatching spoken adjectives and the N400 to incorrect adjectives preceded by matching spoken adjectives [$F(26, 806) = 2.9$, $p = .04$]. These results were confirmed by Bonferroni-corrected post hoc analyses at the Fz electrode on amplitudes during the 350–450 msec window [$t(31) = 2.71$, $p < .01$]. Overall, these data suggest that although the peak amplitude appears to be the same in the three N400 components obtained, they differ in the width of their scalp distribution. Although statistically not significant, a visual inspection suggests that the scalp distribution of the incorrect adjectives preceded by mismatching spoken adjectives is intermediate between the other two subconditions.

The two semantic violations yielded a parieto-central P600 component that appeared to be similar during the first part of that fluctuation regardless of the information contained in Task 2, but differing in a later phase. In this regard, the semantic violation preceded by a mismatching adjective displayed the largest P600 amplitudes. This impression is confirmed by significant main effects of acceptability and electrode by acceptability interactions in all windows after 400 msec, together with significant interactions between matching and acceptability and between matching, acceptability, and electrode in the period from 850 to 1000 msec. Post hoc analyses were performed for the Pz electrode using two wider windows, one for the first part of the P600 (700–850 msec) and a second one for the last part (850–1000).

These analyses confirmed that the P600 did not differ between the two violating written adjectives during the 700–850 msec window [$t(31) = 0.9$, $p > .1$], but did so for the later window [$t(31) = 3.76$, $p < .01$ (all *p*s Bonferroni-corrected)].

Regarding the *syntactic condition*, main results are straightforward and can be summarized as follows: A small LAN and a robust P600 were obtained for syntactically violating adjectives. However, none of these components were affected by the syntactic gender matching/mismatching with the spoken adjective in Task 2. At variance with the semantic condition, no significant effects were found, neither for the matching factor alone nor in interaction with the remaining factors, whereas significant main effects of acceptability and its interaction with electrode cover the whole epoch starting at the 400–450 msec window.

A further finding deserves mentioning. This is the occurrence of an N400 component for all four subconditions within the syntactic condition. If ERPs at parieto-central electrodes are compared across syntactic and semantic conditions, it is apparent that the only deviating subcondition during the N400 interval involves correct written adjectives preceded by semantically matching spoken adjectives in the semantic condition, which displayed a reduced or absent N400. The pervasive N400 component in the syntactic condition is most likely due to the semantic mismatch of all spoken adjectives of Task 2 (see Procedure section). The presence of the N400 in the syntactic conditions was confirmed by planned post hoc comparisons at the Pz electrode for all eight subconditions. Without exception, comparisons between ERPs to correct adjectives preceded by semantically matching spoken adjectives in the semantic condition and the other seven subconditions yielded significant results [$2.5 \leq t(31) \leq 3.8$, Bonferroni-corrected $ps < .01$]. In contrast, no significant difference occurred between these seven subconditions [$-1.6 \leq t(31) \leq 0.2$, $ps > .1$]. This result further reinforces that an N400 can be obtained regardless of the correctness of the critical word by simply introducing semantically mismatching information in Task 2.

DISCUSSION

In separate conditions, the influence of matching or mismatching information held in auditory short-term memory on semantic and syntactic processing of visually presented sentence material was investigated. A main finding of the present study is that syntactic (gender) information within working memory but outside of the currently processed sentence does not influence the syntactic processing of that sentence. No effect was observed either at the behavioral level or on ERP modulations. Both the LAN and the P600 elicited by syntactic violations were identical irrespective of syntactic incongruity, that is, mismatch with spoken adjectives. In addition, no differential modulations were observed for syntactically correct adjectives as a function of the morphosyntactic match with the preceding acoustic material (spoken adjectives). The ineffectiveness of sentence-extraneous syntactic variations is probably not a consequence of any weakness of the syntactic (gender) manipulation as such because it is quite able to produce LANs and rather large P600s, as demonstrated in the present within-sentence violations and in reports by others (e.g., Barber & Carreiras, 2005).

The absence of sentence-extraneous syntactic effects is of interest, considering that morphosyntactic (gender) information as such was held in working memory (Schweppe & Rummer, 2007). Accordingly, the absence of any effect on the syntactic processing of the sentence by external morphosyntactic information would constitute additional evidence for the robust algorithmic nature of the syntactic stream. Nevertheless, further research should determine whether the same result can be obtained when other types of syntactic information are involved. Specifically, word category information should be probed with the paradigm used here because, as proposed by several authors (e.g., Friederici & Weissenborn, 2007), word category may be a more critical syntactic variable for sentence parsing compared to the gender agreement between noun and adjective used here.

Importantly, the opposite observations and conclusions hold for semantic processes. First, acoustic material occupying working memory was able to affect semantic processing at the performance level. Adjectives in Task 2, which were semantically mismatching with adjectives in Task 1, induced more judgment errors, for both correct and incorrect sentences, and facilitated detection of semantic violations. This was the case even despite the different modalities employed for the two tasks, which should have maximized the distinction between sentential (Task 1) and extra-sentential (Task 2) material. Therefore, and at variance with syntactic analyses, semantics would appear at first blush as a heuristic process being vulnerable to external nonsentential influences. Valuable additional information is provided by the ERP fluctuations.

The first striking finding was the presence of an N400 to within-sentence semantically correct material, preceded by a semantically mismatching but sentence-extraneous spoken stimulus. This N400 displayed a centro-parietal topography, latency, and amplitude largely similar to those obtained for semantically incorrect within-sentence adjectives. This finding clearly supports the heuristic nature of the semantic system involved in sentence processing. In this sense, the system has used similar resources as employed when semantic incongruence is detected during sentence processing also for incongruence between sentential and extra-sentential material, which was the case even when the incongruence is between information in different sensory modalities. As mentioned in the Results section, this finding is further supported by the presence of an N400 for all subconditions within the syntactic condition. However, the somewhat wider scalp distribution of the N400 to sentential semantic violations than for correct material affected by sentence-external words indicates that the system does distinguish, at least partially, between incongruence from within and from outside a sentence.

These data contrast with previous reports that the N400 to within-sentence semantic violations and to incongruence between word pairs might be the same (Kutas, 1993). On the other hand, our N400 to correct adjectives was obtained in across-modality word pairing (auditory prime, visual target), a situation that has been found to differentially modulate the N400 (Holcomb, Anderson, & Grainger, 2005; Holcomb & Anderson, 1993). However, cross-modal word pairings have been found to mainly affect the latency and duration of the N400, which clearly has not been the case here. Moreover, our 600-msec SOA between prime and target would seem to facilitate the involvement of the amodal semantic system in which semantic priming presumably takes place (Holcomb & Anderson, 1993). Van Petten (1993) reported a difference between the N400 obtained to semantic violations within a sentence and a semantic mismatch between the critical word and a previous one within the same sentence (sentential context and lexical effects, respectively). The sentential-context N400 was longer in duration and exhibited greater variability across subjects. Although this result shows that an N400 to within-sentence semantic violations and an N400 to a mismatch between words may differ, Van Petten did not report any topographic distinctions, indicating similar mechanisms invoked at different time points. At variance with the experiment of Van Petten, in the present study, the material affecting the processing of the critical word was entirely external to the sentence. Therefore, our results support suggestions of separate semantic subsystems, differing in their specificity for sentence structure (Humphries et al., 2007; Kuperberg, 2007), indicating that at least part of the semantic system may be algorithmic.

Another point of interest about the ERP fluctuations in the semantic condition concerns the P600 component. Interestingly, a P600 was obtained for semantic violations, albeit of lower amplitude as compared to syntactic violations. This result corresponds to our previous findings (Martín-Loeches, Nigbur, Casado, Hohlfeld, & Sommer, 2006) and is in line with a number of recent reports about a P600 to essentially semantic violations, even in syntactically unambiguous sentences (e.g., Kemmerer, Weber-Fox, Price, Zdanczyk, & Way, 2007; Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2007; Kim & Osterhout, 2005; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003). These data would therefore fit with recent proposals about the P600 as reflecting a third, combinatorial linguistic stream. Indeed, our present data would provide evidence that the stream reflected by the P600 is very likely different from the stream reflected in the LAN, as the P600 is sensitive to sentence-extraneous material, whereas this was not the case for the LAN.

According to a current interpretation, the P600 might reflect an independent stream working continuously during sentence comprehension but increasing its activity when a conflict exists either within the syntactic or the semantic streams, or during their integration (Kuperberg, 2007; Kuperberg, Kreher, Goff, McGuire, & David, 2006). In the stream reflected by the P600, both syntactic and semantic information converge in order to achieve the final understanding of the sentence, assigning thematic roles. To the extent that syntactic information is more decisive for these purposes, the P600 to violations of this type would display higher amplitude values than semantic incongruence, but the latter could, nevertheless, elicit a P600 on their own. In view of our results, it appears that this third stream could also be partially heuristic because the P600 to semantic violations, particularly in its later part, was larger if the spoken adjective of Task 2 matched semantically with the violating written adjective. It is as if, in this case, the gender violation was especially detrimental to sentence understanding, an assertion that appears to be reinforced by the particularly widespread scalp distribution of the N400 in this condition. However, two more findings have to be considered here. First, the correct adjective yielding an N400—the one accompanied by a nonmatching spoken adjective—did not elicit a P600. Second, the P600 to syntactic violations was entirely blind to acoustic material. Accordingly, it is possible that the third stream reflected in the P600 is algorithmic in its basic nature, but may be modulated by heuristic processes, at least as outcomes from other streams.

It also appears plausible that this third stream reflected by the P600 is neither syntactic nor semantic in essence, but a combinatorial system exploiting information from different sources with the purpose of achieving final comprehension of the sentence. A recent proposal by Kolk and Chwilla (2007; see also Ye & Zhou, 2008) suggests that the P600 may not even be a specifically linguistic component, reflecting instead an executive control system for alternative interpretations and error checking. Certainly, the P600 has been reported for a number of nonlinguistic tasks, such as harmonic or diatonic violations during the processing of musical information (e.g., Patel, 2003), violations of mathematical sequencing, or mathematical operations with incorrect endings (Martín-Loeches, Casado, Gonzalo, de Heras, & Fernández-Frías, 2006; Nuñez-Peña, Honrubia-Serrano, & Escera, 2005), although in the latter case, and at variance with musical violations, the topography did not perfectly overlap with language-related P600. Anomalies belonging to other nonlinguistic domains, such as geometric patterns (Besson & Macar, 1987) or abstract visual structures (Lelekov-Boissard & Dominey, 2002), have also been reported to yield a P600.

As suggested by some authors (e.g., Münte et al., 1998) and in line with other findings (e.g., Martín-Loeches, Casado, et al., 2006), it is possible that the P600 reflects the composite activity of multiple independent generators, each being responsible for a separate subprocess. This third stream involved in sentence comprehension appears, therefore, as a composite system needing further exploration. Indeed, the differences in topographies, timing, and response characteristics of the LAN and the P600 provide clear evidence that both brain potentials should no longer be supposed to pertain to a single (syntactic) stream. The common claim that such a single system should be encapsulated in a first stage and later open to external sources of information (Friederici & Weissenborn, 2007) further supports the idea of two separate and qualitatively different streams. In our view, the syntactic stream as reflected by the anterior negativities appears entirely algorithmic for sentence parsing. By contrast, the P600 would reflect a different stream for final understanding of the sentence, possibly affected heuristically (or at least by outcomes of heuristic computations in the semantic system) and not exclusively linguistic. Overall, this view would also be compatible with proposals of segregated brain systems for the processes reflected by anterior negativities and the P600 (e.g., Friederici & Kotz, 2003).

Could the observed findings be due to the unavoidable variation in the relationship between the noun and the spoken adjective mentioned in the Methods section? There were no effects of any of the experimental variables during the preadjective epoch, except for a fleeting and short-latency effect of the spoken-adjective–noun relationship. However, one might argue that there might be long-lasting effects beyond 600 msec and reaching into the adjective epoch. As far as the main effect of the matching variable is concerned, such effects were experimentally controlled by balancing because both matching and mismatching conditions contained in equal parts trials were the noun and the spoken adjective had a ''matching'' or ''nonmatching'' relationship. This holds true for both the semantic and syntactic conditions, as can be seen when considering the examples given in the Methods section. However, for the interaction between correctness and matching, a similar balance is not possible. Therefore, the observed interactions in both N400 and P600 between matching and correctness in the semantic condition could be subtly related to such long-latency effects of the relationship between the noun and the spoken adjective. On the other hand, we consider such effects as highly unlikely because they would appear extremely late with—to our knowledge unreported—latencies of around 1000 and 1200–1300 msec, respectively, for the N400 and P600 components.

The noticeable differences between the semantic and the syntactic conditions in RTs and error rates also deserve some comments. The semantic condition could be considered as more difficult, in line with previous results by this group using closely similar materials (Martín-Loeches, Nigbur, et al., 2006), as much as with most previous studies which have manipulated both semantic and syntactic variables within the same experiment (e.g., Hagoort, 2003; Osterhout & Nicol, 1999). Indeed, deciding the syntactic correctness of an adjective relative to a noun in Spanish could be considered a straightforward process because gender is transparent and can be easily checked from orthography. Judging semantic correctness, however, requires accessing world or conceptual knowledge, a sometimes complex task. Nevertheless, these differences should not affect our ERP data, as we essentially performed within-condition comparisons. The only exception has been the comparison of the N400 component across conditions, in support of our finding that an N400 can be obtained even for semantically correct sentence material when preceded by semantically incongruous extra-sentential material. Consequently, as a rule, we measured the effect of extra-sentential material features on trials with similar degrees of difficulty, so that within a given condition certain components related to task difficulty, such as decision-related positivities, should have been similar.

The present dual-task paradigm largely resembles the Reading Span Test for linguistic working memory capacity. In the standard version of this test, subjects must read several sentences for comprehension while simultaneously keeping the last word of each sentence in memory. Despite discrepant opinions about the particular working memory system or subsystem probed by this test (Waters & Caplan, 1996; Daneman & Carpenter, 1980), there is consensus that the last word of each sentence is submitted to working memory. Furthermore, there are diverging views as to whether syntax and semantic information are stored in a language-specific or a general-purpose working memory system (Fedorenko, Gibson, & Rohde, 2006; Waters & Caplan, 1996; Daneman & Carpenter, 1980), and whether syntax and semantic information are stored in the same or different working memory subsystems (Jackendoff, 2007; MacDonald & Christiansen, 2002). Nevertheless, all these proposals agree that information from a given domain (semantic or syntactic) is always stored in one and the same working memory subsystem. Here we studied the effects of syntactic sentence-extraneous information on syntactic sentence processing and similarly for semantic information. Because the essential comparisons were thus always made within a given stream, the interpretation of the present results should not depend on the question of a unitary or multiple working memory system.

In conclusion, our data reinforce a description of the sentence comprehension system as composed of a semantic, a syntactic, and a third integrative stream, devoted to final comprehension. Furthermore, whereas the syntactic system, as reflected by anterior negativities, can still be described as purely algorithmic, both algorithms and heuristics seem to play a role in the semantic system, supporting recent proposals that the latter can be subdivided. The third combinatorial stream, reflected in the P600, is still debated as a system for thematic roles assignment or a general-domain (i.e., nonlinguistic) stream. Our results indicate that both heuristics and algorithms seem to affect this third combinatorial stream.

APPENDIX 1

		Task 1		*Task 2*
Determiner	*Noun*	*Adjective Correct/Incorrect*	*Verb*	*Spoken Adjective Matching/Mismatching*
Semantic Condition				
La	vida	bella/batida	comienza	bonito/agitado
The	life$_{[fem.]}$	lovely$_{[fem.]}$/shaked$_{[fem.]}$	begins	nice$_{[masc.]}$/agitated$_{[masc.]}$
El	perfume	caro/alto	perdura	costosa/crecida
The	perfume$_{[masc.]}$	expensive$_{[masc.]}$/tall$_{[masc.]}$	remains	costly$_{[fem.]}$/large$_{[fem.]}$
Syntactic Condition				
La	fiesta	lujosa/lujoso	empieza	casada/casado
The	Party$_{[fem.]}$	luxurious$_{[fem./masc.]}$	starts	married$_{[fem./masc.]}$
El	Camarero	nervioso/nerviosa	corre	aéreo/aérea
The	waiter$_{[masc.]}$	nervous$_{[masc./fem.]}$	runs	aerial$_{[masc./fem.]}$

Acknowledgments

This research was supported by a grant from the Ministerio de Educación y Ciencia (Acciones Integradas HA2005-0012) and the Deutscher Akademischer Austausch Dienst. We thank Alissa Melinger for careful proofreading of the manuscript.

Reprint requests should be sent to Manuel Martín-Loeches, Center for Human Evolution and Behavior, UCM-ISCIII, Sinesio Delgado, 4, Pabellón 14, 28029 Madrid, Spain, or via e-mail: mmartinloeches@isciii.es; Web: www.ucm.es/info/neurosci/.

REFERENCES

Allen, P. A., Lien, M. C., Murphy, M. D., Sanders, R. E., Judge, K. S., & McCann, R. S. (2002). Age differences in overlapping-task performance: Evidence for efficient parallel processing in older adults. *Psychology and Aging, 17,* 505–519.

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience, 17,* 137–153.

Berwick, R., & Weinberg, A. (1984). *The grammatical basis of linguistic performance*. Cambridge, MA: MIT Press.

Besson, M., & Macar, F. (1987). An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology, 24,* 14–25.

Chwilla, D. J., Brown, C. M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology, 32,* 274–285.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13,* 21–58.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

Fedorenko, E., Gibson, E., & Rohde, D. (2006). The nature of the working memory capacity in sentence comprehension: Evidence against domain specific resources. *Journal of Memory and Language, 54,* 541–553.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47,* 164–203.

Ferreira, F., & Clifton, C., Jr. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348–368.

Ferreira, V. S., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 1187–1199.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 559–585). London: Erlbaum.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6,* 78–84.

Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: Functional imaging and lesion studies. *Neuroimage, 20,* S8–S17.

Friederici, A. D., & Mecklinger, A. (1996). Syntactic parsing as revealed by brain responses: First-pass and second-pass parsing processes. *Journal of Psycholinguistic Research, 25,* 157–176.

Friederici, A. D., & Weissenborn, J. (2007). Mapping sentence form onto meaning: The syntax–semantic interface. *Brain Research, 1146,* 50–58.

Frisch, S., Schlesewsky, M., Saddy, D., & Alpermann, A. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition, 85,* B83–B92.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55,* 468–484.

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience, 15,* 883–899.

Holcomb, P. J., & Anderson, J. (1993). Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language and Cognitive Processes, 8,* 379–411.

Holcomb, P. J., Anderson, J., & Grainger, J. (2005). An electrophysiological study of cross-modal repetition priming. *Psychophysiology, 42,* 493–507.

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2007). Time course of semantic processes during sentence comprehension: An fMRI study. *Neuroimage, 36,* 924–932.

Jackendoff, R. (2007). A parallel architecture perspective on language processing. *Brain Research, 1146,* 2–22.

Johnson-Laird, P. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Kemmerer, D., Weber-Fox, C., Price, K., Zdanczyk, C., & Way, H. (2007). Big brown dog or brown big dog? An electrophysiological study of semantic constrains on pronominal adjective order. *Brain and Language, 100,* 238–256.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52,* 205–225.

Kolk, H., & Chwilla, D. (2007). Late positivities in unusual situations. *Brain and Language, 100,* 257–261.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146,* 23–49.

Kuperberg, G. R., Kreher, D. A., Goff, D., McGuire, P. K., & David, A. S. (2006). Building up linguistic context in schizophrenia: Evidence from self-paced reading. *Neuropsychology, 20,* 442–452.

Kuperberg, G. R., Kreher, D. A., Sitnikova, T., Caplan, D. N., & Holcomb, P. J. (2007). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language, 100,* 223–237.

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research, 17,* 117–129.

Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes, 8,* 533–572.

Kutas, M., & Besson, M. (1999). Electrical signs of language in the brain. In C. Fuchs & S. Roberts (Eds.), *Language diversity and cognitive representations*. Amsterdam: John Benjamins.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203–2005.

Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II (1994–2005). In M. J. Traxler & A. M. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed.). San Diego, CA: Academic Press.

Lelekov-Boissard, T., & Dominey, P. F. (2002). Human brain potentials reveal similar processing of non-linguistic abstract structure and linguistic syntactic structure. *Clinical Neurophysiology, 32,* 72–84.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109,* 35–54.

MacDonald, M. C., & Seidenberg, M. S. (2006). Constraint satisfaction accounts of lexical and sentence comprehension. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 581–612). Oxford: Academic Press.

Marslen-Wilson, W. D., & Tyler, L. K. (1987). Against modularity. In J. L. Gareld (Ed.), *Modularity in knowledge representation and natural-language understanding*. Cambridge, MA: MIT Press.

Martín-Loeches, M., Casado, P., Gonzalo, R., de Heras, L., & Fernández-Frías, C. (2006). Brain potentials to mathematical syntax problems. *Psychophysiology, 43,* 579–591.

Martín-Loeches, M., Nigbur, R., Casado, P., Hohlfeld, A., & Sommer, W. (2006). Semantics prevalence over syntax during sentence processing: A brain potential study of noun–adjective agreement in Spanish. *Brain Research, 1093,* 178–189.

McCann, R. S., Remington, R. W., & Van Selst, M. (2000). A dual-task investigation of automaticity in visual word processing. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1352–1370.

McClelland, J. L., St John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes, 4,* 287–336.

Münte, T. F., Heinze, H. J., Matzke, M., Wieringa, B. M., & Johannes, S. (1998). Brain potentials and syntactic violations revisited: No evidence for specificity of the syntactic positive shift. *Neuropsychologia, 36,* 217–226.

Nuñez-Peña, M. I., Honrubia-Serrano, M. L., & Escera, C. (2005). Problem size effect in additions and subtractions: An event-related potential study. *Neuroscience Letters, 373,* 21–25.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia, 9,* 97–113.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31,* 785–806.

Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes, 14,* 283–317.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6,* 674–681.

Perrin, F., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology, 72,* 184–187.

Reynolds, M., & Besner, D. (2006). Reading aloud is not automatic: Processing capacity is required to generate a phonological code from print. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1303–1323.

Schweppe, J., & Rummer, R. (2007). Shared representations in language processing and verbal short-term memory: The case of grammatical gender. *Journal of Memory and Language, 56,* 336–356.

Sebastián, N. (2000). *LEXESP, Léxico Informatizado del Español*. Barcelona: Ediciones de la Universidad de Barcelona.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

Trueswell, J. C., Tannenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences of parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33,* 285–318.

Van Petten, C. (1993). A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes, 8,* 485–531.

Vissers, C. T. W. M., Chwilla, D. J., & Kolk, H. H. J. (2007). The interplay of heuristics and parsing routines in sentence comprehension: Evidence from ERPs and reaction times. *Biological Psychology, 75,* 8–18.

Waters, G. S., & Caplan, D. (1996). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review, 4,* 761–772.

Ye, Z., & Zhou, X. (2008). Involvement of cognitive control in sentence comprehension: Evidence from ERPs. *Brain Research, 1203,* 103–115.